Epistemic Free Riding

Jeff Dunn
jeffreydunn@depauw.edu

1 Introduction

In epistemology and philosophy of science there has been a growing interest in group inquiry and ways that it might differ fundamentally from individual inquiry. The interest in this topic is understandable. Science is predominately collaborative work. If we want to understand the epistemic success of science, we need to understand group inquiry and it is an important part of this to learn whether it differs from individual inquiry. Philip Kitcher (1990) initiated the focus on scenarios where the norms or best epistemic practices for individuals working alone might come apart from the norms or best epistemic practices for individuals working in groups. Kitcher shows how it is possible for individual scientists who care only about things such as fame and money to nevertheless structure their inquiry in a way that is conducive to finding the truth. Others have carried this work further looking at various ways that group inquiry can differ in important ways from individual inquiry.[1] Uniting much of this work under one heading, Mayo-Wilson *et al.* (2011) have recently investigated different versions of what they call the *Independence Thesis*, roughly, the claim that rational individuals can form irrational groups and that rational groups might be composed of irrational individuals.

In this paper, my goal is to show that some surprising empirical evidence about group problem-solving reveals that groups will often face cases where it is epistemically best for each individual to do one thing, even though this is ultimately epistemically worse for the group. Thus, I will be presenting an epistemic analogue of a free riding scenario. Free riding is familiar in the practical domain, but has seldom been discussed in the epistemic domain, so it is of some interest to investigate whether there are such scenarios. More than that, however, I'll show how the particular epistemic free riding

[1] See, for instance, Strevens (2003), Zollman (2007), Muldoon & Weisberg (2009), Zollman (2010), and Muldoon (2013).

scenario I will discuss directly vindicates a particularly interesting version of the Independence Thesis.

The arguments that I will give here presuppose a kind of epistemic consequentialism. In particular, I'll be taking for granted that what one epistemically ought to believe can be determined by looking at the expected accuracy of believing in that way.[2] Below I discuss in more detail the particular form of epistemic consequentialism my arguments rely on, but I don't defend the consequentialist viewpoint here. Rather, this paper should be seen as illustrating one interesting consequence of adopting the consequentialist picture.

Here's the plan. In section 2 I'll give a precise characterization of free riding scenarios. After this, in section 3, I'll consider the only other example in the literature (of which I'm aware) of epistemic free riding and explain why it doesn't actually seem to be a case of epistemic free riding. In section 4 I'll present the empirical evidence that I'll argue leads to an epistemic free riding scenario. In section 5 I'll give a formal model to rigorously show that the scenario I'm considering does constitute a free riding problem. In section 6, I'll consider the wider significance of the existence of epistemic free riding and how it relates to the Independence Thesis. Finally, in section 7, I'll explain how the epistemic free riding scenario gives rise to an interesting puzzle about the explanation for some of our natural doxastic dispositions in cases of group inquiry.

2 Characterizing Free Riding Scenarios

What characterizes a free riding scenario? I will follow Philip Pettit (1986) in maintaining that a free riding scenario arises when two conditions are met.[3] First, there is some behavior such that if everyone behaved in this way, the outcome would be Pareto-inferior to the outcome where everyone does not behave in that way.[4] Call this the *Pareto condition*. Second, that behavior can nevertheless be justified in terms of self-interest: the free rider really stands to gain from behaving as he does. Call this the *self-interest condition*. These two features together are what make free riding scenarios

[2]As I'll note, there are ways of running the argument here that don't take accuracy as the sole epistemic good. However, any way of running the argument relies on some form of consequentialism.

[3]Pettit goes on in that article to distinguish between two types of scenarios where both conditions are met. He dubs one type of scenario as giving rise to *free riding* and the other to *foul dealing*. I won't make those finer distinctions here.

[4]An outcome, O1, is Pareto-inferior to outcome, O2, iff in O1 no one is better off and at least some are worse off compared to O2.

especially pernicious. The satisfaction of the Pareto condition shows that the behavior is of the sort that is damaging to everyone if everyone engages in it, and yet the satisfaction of the self-interest condition shows that everyone can justify behaving in that way on the basis of self-interest.

In many presentations of practical free riding, a dominance argument is given to show that the self-interest condition is met. A dominance argument for some action is one that demonstrates that, no matter what the state of the world is, it is best for you to perform that action. The argument for why it is your self-interest to defect in a prisoner's dilemma is a dominance argument: no matter what your partner does, you always do better if you defect. More formally, if C_i (D_i) is the strategy of player i cooperating (defecting), and if '$>_i$' means 'is preferred by i', we have: $(D_1, C_2) >_1 (C_1, C_2)$ and $(D_1, D_2) >_1 (C_1, D_2)$, where you are player 1 and your partner is player 2.

Inspection of many standard examples of free riding, however, reveals that the justification for free riding behavior is not always a strict dominance argument. Suppose that there is a group of recreational fishermen who work a small lake and to prevent overfishing, a quota for the season has been set for each person. If everyone sticks to his quota, this will ensure that there is enough fish to have productive fishing seasons year after year. However, there is a safety buffer built into the quotas so that if everyone catches his quota, this does not put the fish population on the brink of disaster. Suppose, then, that everyone adheres to his quota. Then, it is in my interest to ignore the quota, for if I alone exceed my quota, this will not deplete the fish population to dangerous levels. I ensure myself a bigger catch than others this year while still ensuring a replenished supply of fish for next season. Suppose that no one is sticking to his quota. Then, it is again in my interest to ignore the quota, for if everyone is exceeding the quota the population will be decimated whether I stick to it or not. I may as well get as many fish while I can. What about the intermediate situations, where some but not all are sticking to the quota? It is hard to see why it is not in my interest to ignore the quota here, too. For, on the one hand, suppose that there are enough fishermen sticking to the quota that the fish population will survive until next year. In that case, it is unlikely that *my* taking a few more fish will tip the balance. On the other hand, suppose that there are enough fishermen ignoring the quota that the fish population will not survive. In that case, *my* sticking to the quota isn't going to help and so I may as well ignore it. Thus, it seems that no matter the state of the world, it is best for me if I ignore the quota. This seems to be a dominance argument to justify free riding.

On reflection, however, it isn't actually a dominance argument. For surely there is some amount of fish such that if less than that amount is caught then the fish population in the small lake survives and if more than that amount is caught then the fish population does not survive. Given this, there is some state of the world where *my* decision about whether to ignore the quota or not is the decision that determines whether or not the fish population survives. And if my decision to ignore the quota is the one that decimates the fish population for years to come, then it is not in my interest to ignore the quota in that state of the world.[5] Thus, we don't have a dominance argument for free riding. Nevertheless, the reasoning in the previous paragraph still provides a compelling reason for me to free ride. The reason for this is that it is extremely difficult to tell whether I am the crucial one to push the fish population to dangerous levels. Further, if there are a large number of other fishermen, it is extremely unlikely that I am in this situation. This suggests that a free riding scenario can obtain even when there is no *dominance* argument for free riding behavior. The free rider need not prefer the outcome where he defects to the outcome where he cooperates no matter the state of the world. What matters, instead, is that the free rider must be able to justify free riding behavior in terms of self-interest. A dominance argument is a particularly clear way to do this, but it is not the only way.[6]

The example of the fishermen is a practical free riding scenario: the value is prudential and the things being evaluated are actions. In what follows, I'll give an example of epistemic free riding. An epistemic free riding scenario is formally identical to a practical one. The key difference is that the value is *epistemic* and the things being evaluated are *beliefs*.

In some sense, it is extremely easy to construct an epistemic free riding scenario. Suppose that there are 100 of us in a group and if over half of us believe that the earth is flat, then an oracle will tell us 1,000 important truths about the world. If accuracy of belief is the only thing of epistemic

---

[5]This can be put more formally using the same notation as above. Suppose there are 5 fishermen. What the argument in the text suggests is that while

$(D_1, C_2, C_3, C_4, C_5) >_1 (C_1, C_2, C_3, C_4, C_5)$

and while

$(D_1, D_2, D_3, D_4, D_5) >_1 (C_1, D_2, D_3, D_4, D_5)$

there may be some specific ways that the other fishermen can defect or cooperate where the following holds:

$(D_1, D_2, D_3, C_4, C_5) <_1 (C_1, D_2, D_3, C_4, C_5)$.

This will be the case, for instance, if the fish don't survive the next season in $(D_1, D_2, D_3, C_4, C_5)$ but do survive in $(C_1, D_2, D_3, C_4, C_5)$.

[6]For some thoughts on this, see Pettit (1986), p. 369.

value, then failing to believe that the earth is flat seems to be an instance of epistemic free riding. If you fail to believe the earth is flat, you stand to have 1,000 true beliefs and no false beliefs by free riding on the fact that others in the group will believe the earth is flat. This meets the formal conditions for a free riding problem and it is recognizably epistemic. However, there is not much interest in such a case, because the conditions for it to obtain are so unrealistic. Practical free riding is interesting precisely because the conditions for it to occur often arise. Analogously, epistemic free riding will be interesting only if the conditions for it to occur often arise. This toy example does not show us *that*. In contrast, I will present a case of epistemic free riding that is realistic and often arises. Before getting to such a case, however, we need to look at the only other example in the literature of epistemic free riding.

3 List & Pettit on Epistemic Free Riding

Christian List & Philip Pettit (2004) purport to give us an example of epistemic free riding that, unlike the oracle scenario, is more realistic. Their scenario makes use of the Condorcet Jury Theorem (CJT). Consider a case where a group is trying to decide the answer to a question with two possible answers (e.g., guilty or not guilty). The CJT says that if each member is more likely than not to choose the correct answer, and if each member decides independently of the other members, then the probability that the majority vote is accurate approaches 1 as the number of group members increase.[7] This initially striking mathematical result points toward one way that groups can sometimes be more accurate than any individual member of the group.

So, how can this lead to a free riding scenario? List & Pettit ask us to imagine a situation where we are taking votes sequentially and in public. Further, we are to assume that we all know that we are each equally likely to get the verdict correct, that this probability is greater than 0.5, and that we believe each person is voting in accordance with his or her belief. In this case, and if we each know about the CJT result, then we each know that it is more likely that the majority vote is correct than any one member is correct. Thus, if I am voting third and I have heard two *yea* votes, then no matter

[7] It ends up being important exactly how these different conditions are formally specified to understand the exact content of the theorem. Further, there are extra conditions one can add to strengthen the result. However, these details won't matter for our purposes. A good source for more information on CJTs is Hawthorne (ms).

what I independently believe to be correct, I stand a better chance of voting correctly if I vote *yea*. And the same thing is true for each person after me. However, in so-doing, we each make the group significantly less likely to get the right answer, since it is now a group with only two independently voting members.[8]

An important question is whether this scenario really is a case of epistemic free riding. As List & Pettit present the example, the members of the group value voting accurately themselves. Since the value in question concerns accurately answering a question, we are plausibly in the epistemic domain. But it is doubtful that it meets the conditions for a free riding scenario. The self-interest condition states that there must be an argument for why it is in the self-interest of the free rider to free ride. Note first that this scenario is not one where there is a dominance argument for voting non-independently. Suppose that *yea* is the correct answer but that, improbably, a large majority is voting *nay*. If agent 1 values voting correctly herself we have: $(N_1, N_2, \ldots, N_n) <_1 (Y_1, N_2, \ldots, N_n)$. This shows that there is no dominance argument for voting with the group. Rather, it is some sort of expected value calculation that must be doing work. But, as noted by List & Pettit, such an argument works only if we assume that everyone else is voting independently. But if one is reasonably confident that everyone else is free riding and voting non-independently, then one has no reason to go with the group since the majority vote is no longer more likely to be correct than one's own.

The scenario also fails to satisfy the Pareto condition. By the very setup of the case, it is not possible for everyone to engage in free riding behavior, which in this case is to simply vote the way the group votes rather than independently. For the first voter, this is impossible. So, in this scenario, we cannot even evaluate whether it is true that everyone engaging in free riding would be Pareto-inferior to everyone not free riding.[9] Thus, this

[8] If the first two votes are split, then the same situation obtains for the person who votes fifth so long as the next two votes agree with each other. In the economics literature, these are known as *informational cascades*. For seminal work in this area see Banerjee (1992) and Bikhchandani *et al.* (1992).

[9] One might try to defend List & Pettit by slightly weakening the Pareto condition. The Pareto condition currently says: Behavior b meets the Pareto condition iff the outcome that results from **everyone** choosing b is Pareto inferior to the outcome where no one chooses b. One might weaken it to say instead: Behavior b meets the Pareto* condition iff the outcome that results from **most** choosing b is Pareto inferior to the outcome where no one chooses b. One problem with the Pareto* condition is that it is vague: how many in a group must choose b for it to be the case that most of the group members chose b? But perhaps a modification of the Pareto condition along these lines could be made to

scenario—though interesting in many ways—isn't an instance of epistemic free riding.

List & Pettit have an alternative scenario, however. In this alternative scenario the members of a group only care that the majority vote is accurate. This is the sole outcome with epistemic value. It's not altogether clear why the group getting the correct answer is of any epistemic value to me, a group member. But put this aside. Perhaps we can make this assumption plausible by assuming that group members will adopt as their own belief whatever the majority votes for. Each member, then, wants the majority vote to be accurate, but each member would also like to save herself the time-consuming effort of coming to an independent view on the issue. It looks, at least initially, as if voting non-independently is an instance of free riding. First, it seems as though the self-interest condition is satisfied. For, ignoring the case where my vote is pivotal and pushes the majority vote to one side or the other, I can save myself significant effort by voting non-independently without changing how the group votes. If we're going to get the same answer regardless of my vote, I might as well save myself the effort of thinking through the issue. And second, the Pareto condition might appear to be satisfied, since we're a much more accurate group when we all vote independently than when we do not.

As with the scenario above, this is an interesting case. However, I also doubt that this is a case of epistemic free riding. The first reason is familiar: initial appearances to the contrary, the Pareto condition cannot really be satisfied in this case, since it is impossible for everyone to vote non-independently. To vote with the group, rather than based on an independent analysis of the issue requires that the group is voting some way and it is hard to see how this can happen unless at least one person votes independently.

The second worry is novel, however. The worry is that the argument that non-independent voting is in my self-interest concerns more than just epistemic value. The argument also involves prudential value like the time and effort it takes to come to an independent answer. For this reason, the scenario is not a case of *epistemic* free riding.

I do not want to discount the interesting nature of these two scenarios presented by List & Pettit. They certainly bear a family resemblance to classical free riding problems, they in some way concern epistemic matters, and they strike me as interesting in their own right. Nevertheless, I doubt that we have been given a realistic scenario that truly can be said to be a

work. Still, the scenario would fail the self-interest condition.

case of epistemic free riding. This is important especially if we want to use a free riding scenario to vindicate an *epistemic* version of the Independence Thesis. In what follows, I aim to present such a case. It is, I claim, a realistic scenario, it is purely about epistemic value (not prudential value), and it meets both the Pareto and self-interest conditions.

4 Empirical Evidence

To get to the scenario I want to focus on, we need to briefly consider some psychological research on group inquiry. There is now considerable evidence for the claim that groups that have members who debate a question *and* genuinely hold dissenting views about the debated question (that is, they are not simply playing the devil's advocate) are more accurate than both individual inquirers and other groups that lack one of these qualities. One important consequence of this is that heterogenous groups—groups where there are individuals that hold different views about the correct answer to a question—are more likely to eventually reach the correct answer to the question than homogenous groups—groups where everyone shares the same view on the answer.

The evidence for this surprising claim comes from several sources. First, there is evidence that groups can often outperform any individual member in reasoning tasks. Consider a study by Moshman & Geil (1998). In the study, the participants were divided into 3 experimental conditions: individual control, interactive condition, individual/interactive condition. In the individual control, the participants were asked to solve the Wason Selection Task on their own. In the interactive condition, participants were asked to solve the task in groups with 5-6 members. In the individual/interactive condition, participants were first asked to solve the task alone, and then (without having the correct answer revealed to them) solve the task in a group. The results are striking. In the individual control condition, consistent with other studies on the Wason Selection Task, the success rate was approximately 9%. In the interactive condition, the success rate jumped to 70%. Finally, in the individual/interactive condition, when these individuals worked in *groups*, the success rate was 80%. So, groups can often be more accurate than individuals working alone.

However, there is an important caveat to this general result. There is evidence that for groups to do better than individuals, group members must be debating and arguing with each other in a genuine way. As Mercier & Sperber (2011) write: "...many experiments have shown that *debates* are essential to any improvement of performance in group settings." (p. 63, my emphasis). Schulz-Hardt *et al.* (2006) is a good example of the kind of experiment demonstrating this. The study involved groups attempting to solve *hidden profile* problems. These are problems where the correct solution requires full information, but no one group member possesses full information. For instance, the problem might be to select the best apartment. In a hidden profile problem, the full set of information clearly picks out one apartment as best, but each group member only has partial information. Groups discuss the problem and then decide on a group answer. Schulz-Hardt *et al.* manipulated the information available to each group member to test the effect of diversity of opinion on the eventual solution the groups put forward. Groups where all group members had full information reached the correct decision 100% of the time. Groups with homogenous preferences for options before discussion reached the correct decision 7% of the time. Groups where group members had diverse preferences before discussion—but no member who initially preferred the correct option—reached the correct decision 26% of the time. Finally, groups where group members had diverse preferences before discussion—and at least one member who initially preferred the correct option—reached the correct decision 62% of the time. Overall, groups with heterogenous views reached the correct decision 43% of the time.[10]

In an early review article of the literature on group problem-solving, Hastie (1986, pp. 151-2) identifies three characteristics that produce high levels of group performance. The first is that the individuals vary in their competency to answer the problem, and where the problem has a "eureka solution", a solution that may not be obvious initially but is demonstrable once discovered. Second, individual judgment accuracy is perturbed by unsystematic errors. This characteristic allows that simple averaging of group members' answers is more likely to be correct than any individual judgment. The third characteristic that leads to high level group performance is group members who possess different evidence. It is plausible that some combination of these characteristics often obtain in cases that we care about. The first and third conditions are likely to obtain in certain scientific settings, for instance, when a group of physicists is discussing the proper interpretation of some data from the Large Hadron Collider. Though the correct interpre-

[10] See also Perret-Clermont *et al.* (2004) (who focus on childhood development), Schulz-Hardt *et al.* (2006) (for a helpful summary of this research on adults), and Kuhn *et al.* (1997). Other evidence for the claim that debate is needed for groups to outperform individuals is summarized in Mercier & Sperber (2011); Mercier (2012, 2011). The claim also fits well with what Sunstein (2002) calls the "law of group polarization". For skepticism about the robustness of this phenomenon, see Gigone & Hastie (1997).

tation is by no means *simple*, it is demonstrable, once found. Further, it is likely that different physicists bring slightly different bodies of evidence to bear on the problem. The first and second conditions are plausibly met in the key cases used to motivate conciliatory views of peer disagreement in the recent epistemology literature, such as a case where a group of us disagree about our share of the check at a restaurant (Christensen, 2007).

The data so far canvassed suggest that groups of inquirers that engage in genuine debate about possible answers to a question are ultimately more accurate in answering the question than groups of inquirers that do not engage in such debate. But the results are more surprising than this. It turns out that contrived debate using various devil's advocacy techniques does not yield the advantages for group inquiry that genuine debate among group members who genuinely hold dissenting views. Schulz-Hardt *et al.* (2002), for instance, investigated whether various groups that consisted of genuinely disagreeing members perform similarly to groups where dissent is contrived using various devil's advocacy techniques. They found that groups that had genuine disagreement were less biased in seeking only confirmatory information. Greitemeyer *et al.* (2006) set out to show that, contrary to the results just described, contrived dissent *can* yield advantages for group reasoning. They instructed various group members to defend different points of view, even if that point of view was not one the group member really believed. Though group members did play the appropriate roles, and as a result a more balanced menu of evidence was discussed, the group answers did not improve. In contrast, in the genuinely heterogenous groups, group answers did improve markedly.[11]

Putting all this evidence together provides a strong reason to think that, at least in many kinds of cases, group members who are investigating some question will, in the long run, end up with more accurate beliefs about the answer to that question if they initially maintain their divergent beliefs and vigorously defend them. Of course, if the group members are to reap the benefits of this group discussion, they must eventually converge on one answer. But they are more likely to believe the true answer if, initially, they stick to the answer they think is true in spite of the disagreement with their

[11] Strauss *et al.* (2011) summarizes this study as well as other related ones. Indirect support for the claim that genuine dissent is required for the advantages of group inquiry is provided by the well-known phenomenon of belief bias whereby people are better at identifying flaws in arguments when the conclusions are those with which they disagree and worse at identifying flaws in arguments when the conclusions are those with which they agree (Evans *et al.*, 1983).

peers.[12]

5 Epistemic Free Riding

From this empirical data we can construct an epistemic free riding scenario. Suppose that I am in a relatively large group of experts on some topic attempting to answer some question in that topic. Before we start discussing the question, we realize that opinion is approximately evenly split between the possible answers. Plausibly, once the fact that there is this split of opinion among experts is added to my evidence, my evidence now supports each answer to a roughly equal degree. So, if I want to respect this evidence, I should reduce my confidence in my preferred answer and adopt a roughly similar credence for all the possible answers being entertained. However, if all the group members do this, then we will all go into the debate with the same agnostic belief state with respect to our question. We will have turned ourselves from a heterogenous to a homogenous group, thus severely reducing the chance that we will reach the correct answer to the question at the end of debate. Thus, if we are looking only at epistemic value, we'd each prefer to be in a group containing mostly members who remain steadfast in their initial opinions during the debate. This way we maximize our chance of landing on an accurate consensus answer to our question. However, since our evidence now supports withholding belief, we each have a good reason to now withhold belief and not remain steadfast in our pre-debate opinions.

That's the basic scenario, but to really make good on the claim that it is a case of epistemic free riding, it needs to be made more precise. To make it more precise we need to specify several things. First, we need to say something about epistemic value. Second, we need to say something about epistemic acts. And third, we need to say something about how epistemic acts are evaluated with respect to epistemic value.

As to the first point, I'll assume that the only thing of epistemic value is accuracy.[13] I'll also assume that accuracy comes in degrees. For example, although a credence of 0.5 in P and a credence of 0 in P are both inaccurate

[12] There's an obvious potential challenge here for conciliatory views of disagreement, views according to which (roughly) one ought to suspend judgment on a proposition disputed by one's peers. This challenge is considered in Dunn (ms).

[13] The argument doesn't require this; one could still run the argument if one thinks there is, say, epistemic value in following the evidence, and epistemic value in knowing. Taking accuracy as the only epistemic value, however, allows the argument to be continuous with other formal work that appeals to scoring rules (e.g., Joyce (1998) and the work that follows in this tradition).

when P is true, a credence of 0.5 in P is more accurate than a credence of 0. To measure this precisely we can use a scoring rule. Let the question of interest have a finite number m of mutually exclusive and jointly exhaustive possible answers in $\Omega = \{1, \ldots, m\}$. Let a belief state with respect to this question $\mathbf{c}$, be a probability vector $\{c_1, \ldots, c_m\}$ such that $c_1, \ldots, c_m \geq 0, c_1 + \ldots + c_m = 1$. So, for instance, c_1 is the credence assigned to answer 1, c_2 the credence assigned to answer 2, etc. If $\mathcal{P}_m$ is the set of all probability vectors of length m, then a scoring rule is a function $S(\mathbf{c}, i) : \mathcal{P}_m \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$. Scoring rules thus award scores to credences based on only two things: the level of credence assigned to the answer and the state of the world. Thus, they are measures of *accuracy*.

There are many scoring rules that can be used. I'll use the popular Brier score in the text (and introduce a second one in the Appendix): The Brier score is given by:

$$B(\mathbf{c}, i) = \sum_{j=1}^{m} (\delta_{ij} - c_j)^2,$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. Note that the Brier score has a minimum value of zero and a maximum value that increases with increasing m. If it is important to compare scores between questions that have different numbers of possible answers we can normalize the Brier score by taking $B(\mathbf{c}, i)/m$ where m is the number of possible answers to the question. According to the Brier Score, lower numbers are better. Thus, we can see the Brier score as measuring inaccuracy: less is better.

That covers epistemic value. As for epistemic acts, I'll suppose that agents can perform epistemic acts at times and that the only epistemic acts available to an agent at a time are the belief states that the agent can come to occupy at that time.[14] Since in this context we'll be focusing on specific questions that groups are attempting to answer, we'll focus solely on the acts available to the agent that concern her belief state *with respect to the question of interest*.

Finally, epistemic acts will be evaluated in virtue of their *expected* accuracy. Why *expected* accuracy rather than just accuracy? There are several reasons, but the easiest way to see why expected accuracy is needed is by noting that a proposition could be true even if a tremendous amount of evidence points towards its falsity. In such a situation, it seems it is epistemically best to believe the proposition is false, even though such a belief is inaccurate.

[14] Greaves (2013) makes this a condition of something being an epistemic act.

Given this set up, we can now think more specifically about the scenario of group inquiry. To keep things as simple as possible, we'll consider a scenario where a group of experts is interested in answering a question that has only two possible answers: $\{P, \bar{P}\}$.[15] We'll also suppose that there are only three belief states that an agent can adopt: $\{c_P = 1, c_{\bar{P}} = 0\}$, $\{c_P = 0, c_{\bar{P}} = 1\}$, and $\{c_P = 0.5, c_{\bar{P}} = 0.5\}$. These correspond, respectively, to believing P, believing $\bar{P}$, and withholding belief with respect to P. Denote these with '$\mathbf{c}^P$', '$\mathbf{c}^{\bar{P}}$', '$\mathbf{c}^W$'. We suppose that initially our group contains about equal numbers of members who believe P and who believe $\bar{P}$. So, initially, each group member starts out opinionated. There are then two key points at which we'll evaluate our agents' expected accuracy. First, upon realizing there is a disagreement, each agent will either stick with her antecedent view and remain opinionated or she will withhold belief. Call these options, respectively, the option of being *steadfast* and of being *conciliatory*. The group will then discuss and debate the issue and then adopt some consensus answer to the question. We then come to the second evaluation point. We'll assume that each agent adopts the group's consensus answer as her own at the end of the debate. Obviously, each agent's initial decision about whether to be steadfast or conciliatory has an effect on her initial expected accuracy. But since her decision here can also in part affect whether she is in a heterogenous or homogenous group, it can also affect how likely it is that the group consensus is correct.

With an eye toward whether being conciliatory is to be an epistemic free rider, let's consider the kinds of groups that some agent, S, could be in. First, S could be in a group where everyone is conciliatory. In that case, it is plausible that S being steadfast doesn't make her group have the required heterogeneity to get the benefits of group debate. Second, S could be in a group where she is the pivotal member in the sense that if S is steadfast, then the group will have the requisite heterogeneity to get the benefits of debate, but if S is conciliatory then they will not. Third, and finally, S could be in a group that has enough other members who are steadfast that the group will get the benefit of heterogeneity independent of what S does. As in the classic free riding scenarios discussed above, we ignore the pivotal case. This yields the following decision matrix for S (where a, b, c, and d are the epistemic payoffs for S):

[15] I'll let '$\bar{P}$' abbreviate the negation of P.

	All other group members are conciliatory	Enough other group members are steadfast
S is conciliatory	a	b
S is steadfast	c	d

I maintain that the epistemic act of being conciliatory is often an act of free riding. This is so just in case the self-interest condition is met as well as the Pareto condition. Given the way we've set things up, the self-interest condition is met so long as $a < c$ and $b < d$ (recall that these are expected *inaccuracy* scores and so lower values are better). We can see whether the Pareto condition is met by determining whether $d < a$.[16]

Let C and S correspond to S's initial epistemic acts of either being conciliatory or being steadfast. Let GC and GS correspond to the epistemic act of adopting the group consensus when S is in either a conciliatory or a steadfast group. Letting $EB(X)$ denote the expected Brier inaccuracy of epistemic act X, we have:

$$a = EB(C) + EB(GC)$$

$$b = EB(C) + EB(GS)$$

$$c = EB(S) + EB(GC)$$

$$d = EB(S) + EB(GS)$$

To see if the Pareto and self-interest conditions obtain, we need to say something about $EB(X)$, the expected value of taking epistemic act X. Let i range over the possible answers to the question (in our case: P and $\bar{P}$). Let $\mathbf{c}^n$ range over the possible belief states that can be adopted (in our case: '$\mathbf{c}^P$', '$\mathbf{c}^{\bar{P}}$', '$\mathbf{c}^W$'.). The acts and possible belief states can each be subscripted with a time to indicate at what time the act is an option. Finally, let p be a probability function that tracks what the agent's evidence supports now. Then, we have:

$$EB(X_t) = \sum_i \sum_n p(i \wedge \mathbf{c}^n_t | X_t) \times B(\mathbf{c}^n_t, i)$$

Several points deserve note. First, the weight for this expectation is a probability function, p, that tracks what the agent's evidence supports, but which

[16]Roughly, the Pareto condition is met iff the case where all free ride is worse than the case where all do not free ride. One might notice, however, that outcome d is not necessarily a case where *all* do not free ride; it is instead a scenario where *some* do not free ride. However, as will be clear shortly, the value of d in this model will be the same in the case where all are steadfast and in the case where enough are steadfast.

need not be identical to the agent's own belief state. There are several reasons for this. First, there is a technical reason: the agent's belief state with respect to the question, **c**, is only defined over possible answers to the question. So it doesn't have enough structure to tell us, say, how likely it is that the group consensus is accurate given that the group members hold heterogenous views. Second, there is a philosophical reason: we are interested in whether there might be situations where what an agent should believe (for consequentialist reasons) comes apart from what the evidence supports. So, it is useful to distinguish what the evidence supports from what the agent believes. Finally, another philosophical reason: one way to maximize expected value (of any kind) is to maximize relative to how one's own degrees of belief weight the different possibilities; but another way to maximize expected value is to maximize relative to more objective weights that track what your evidence supports. It is this latter way that we are interested in here.[17] So, we are asking what is going to maximize epistemic value for an agent, given what her evidence supports.

The second thing to note about $EB(X)$ is that it does not represent a fully general way to calculate the expected epistemic value of adopting a belief state. It builds in, for instance, that we are evaluating act X only with respect to the expected accuracy that X has now rather than with respect to the expected accuracy into the future. This is acceptable in this context since by ignoring the case where the agent is pivotal to whether the group is heterogenous or homogenous, her initial decision doesn't have an effect on her later decision to adopt the group consensus.

Finally, by using a conditional probability, $EB(X)$ also takes a stand in the dispute between evidential and causal decision theory, preferring evidential decision theory. This feature might be objectionable, but in light what was said in the previous paragraph, it won't be a problem in this case.[18]

Let us now consider whether the self interest condition is satisfied. To do this, we need to see whether $a < c$ and $b < d$. The second term in each of these inequalities is the same, so our question reduces to whether $EB(C) < EB(S)$.

So, we'd like to work out (I've dropped the time subscript for ease of reading):

$$EB(S) = \sum_i \sum_n p(i \wedge \mathbf{c}^n | S) \times B(\mathbf{c}^n, i)$$

[17]Ralph Wedgewood (ms) argues for a similar setup.

[18]For more on the distinction between evidential and causal decision theory in the epistemic realm, see Greaves (2013) and Konek & Levinstein (ms).

$$EB(C) = \sum_i \sum_n p(i \wedge \mathbf{c}^n | C) \times B(\mathbf{c}^n, i)$$

These can be simplified. Since choosing act C at the beginning of inquiry will result in having belief state $\mathbf{c}^W$ at the beginning of inquiry, it follows that $p(\mathbf{c}^W|C) = 1$ and so $p(i \wedge \mathbf{c}^W|C) = p(i|C)$ and $p(i \wedge \mathbf{c}^P|C) = p(i \wedge \mathbf{c}^{\bar{P}}|C) = 0$. Further it is plausible that my choosing to be conciliatory does not affect the probability that answer i is true. Thus, $p(i|C) = p(i)$. Accordingly, $EB(C)$ can be simplified to:

$$EB(C) = \sum_i p(i) \times B(\mathbf{c}^W, i),$$

where i can take value either P or $\bar{P}$.

A very similar line of argument simplifies $EB(S)$ to either of the following, depending on whether the agent in question either initially believed P or $\bar{P}$:

$$EB(S) = \sum_i p(i) \times S(\mathbf{c}^P, i)$$

$$EB(S) = \sum_i p(i) \times S(\mathbf{c}^{\bar{P}}, i).$$

If our scoring rule is symmetric, as the Brier score is, in that it assigns the same penalty to $c(P) = n$ when P is true as it assigns to $c(P) = 1 - n$ when P is false, then these two expected scores will be equal.

So long as we use a proper scoring rule[19], as the Brier score is, these expectations will be minimized for $c_i = p(i)$. So, what is the value of $p(P)$ (and thus also $p(\bar{P})$) in this situation? It is plausible to hold that it is at or near 0.5. This is because at the time of the decision we have a number of qualified individuals, half of whom believe P and the other half who believe $\bar{P}$. The conciliatory view in the peer disagreement literature contains arguments that support this claim.[20] With such a value for $p(P)$, one minimizes expected inaccuracy by adopting $\mathbf{c}^W$, which assigns P a credence of 0.5 and $\bar{P}$ a credence of 0.5. Accordingly, $EB(C) < EB(S)$. From this it follows that the self-interest condition is met.

Turn now to the Pareto condition. Here we want to know whether $d < a$, that is, whether $EB(S) + EB(GS) < EB(C) + EB(GC)$. From what was

[19]A proper scoring rule is a scoring rule for credences that has the following property: the credence function that has the best expected score from the perspective of any coherent credence function, c, is c itself. For more on this see Seidenfeld (1985).

[20]This view is sometimes called the 'Equal Weight View' (Elga, 2007).

just shown we know that $EB(C) < EB(S)$. The question, then, is whether $EB(GS) < EB(GC)$ to a degree great enough to make $d < a$. To answer this question we need to pay closer attention to our scoring rule. If we normalize the Brier score, $EB(C) = 0.25$ and $EB(S) = 0.5$. Thus, we want to know under what conditions, $EB(GS) + 0.25 < EB(GC)$.

Let 't_{GS}', 'f_{GS}', and 'w_{GS}' denote the probability that a group reaches a true answer, a false answer, or withholds belief conditional on the group being steadfast (and similarly with 'GC' replacing 'GS'). Note that there are two ways to adopt a true (false) belief after the debate: (1) adopt $\mathbf{c}^P$ when P is true (false) or (2) adopt $\mathbf{c}^{\bar{P}}$ when P is false (true). However, as noted, the Brier score is symmetric and so yields the same score or each of these cases. In particular, the score for a true belief is 0, the score for a false belief is 1, and the score for withholding belief is 0.25. Since at the end of inquiry the agent is going to adopt the group consensus, we know that $EB(GS) = t_{GS} \times 0 + f_{GS} \times 1 + w_{GS} \times 0.25$. From this it follows that $d < a$ if and only if:

$$4f_{GC} + w_{GC} - 1 > 4f_{GS} + w_{GS}.$$

If groups do not withhold belief at the end of inquiry, then this holds so long as $f_{GC} - 1/4 > f_{GS}$. One simple scenario where this obtains is where $f_{GC} = t_{GC} = 0.5$ and where $f_{GS} = 0.24$ and $t_{GS} = 0.76$. But even if we allow groups to withhold belief at the end of inquiry, there is still a wide range of cases where the inequality holds and so the Pareto condition is met. For instance, here is such a scenario:

$$t_{GC} = 0.4 \quad w_{GC} = 0.2 \quad f_{GC} = 0.4$$

$$t_{GS} = 0.7 \quad w_{GS} = 0.15 \quad f_{GS} = 0.15$$

Since we have good empirical reason to believe that diverse heterogenous groups are often much more likely to be accurate than homogenous groups, we have reason to believe that in actual cases of group inquiry, both the Pareto and self-interest conditions can be met. Thus, the option of being conciliatory and withholding belief upon realizing that there is a mix of opinion on a question is an epistemic act of free riding.

Stripping off the formalism, here is the basic picture. In cases of group inquiry, think of group members as able to garner epistemic value at two times: at the beginning of inquiry and then at the end. Once a split of opinion is noticed, each group member expects to be most accurate at the beginning by withholding belief on the answer to the disputed question. Each group member also plans to adopt the consensus answer at the end of inquiry. But no one group member can do much to affect the consensus answer at the later time, so since each group member can get a boost in epistemic value *now* by withholding belief, each is reasonable in withholding belief. However, if each follows this advice then the group will be homogenous in their opinions and is less likely than they otherwise would be to reach an accurate consensus answer at the end of inquiry. So although each group member prefers that she withholds belief at the beginning of inquiry, she hopes everyone else will not.

This is surprising in some ways. In the practical domain, actions that count as free riding actions are usually thought of negatively. Not paying one's taxes or fishing above a quota are classic examples of free riding. In this epistemic case, however, the free riding action is one that is often looked on in a favorable light. Someone who modestly withholds belief once she realizes that there is a dispute amongst her peers is generally thought to exhibit some sort of intellectual virtue. The model here together with the empirical evidence suggests that it is nevertheless an act of epistemic free riding.

Before closing this section, let me note something about several simplifying idealizations made in the model. First, we considered a question with only two possible answers. However, the model is able to handle scenarios where questions have more than two possible answers. In the Schulz-Hardt *et al.* (2006) study cited above, the experimental data is gathered for groups solving a question with four possible answers (in particular, the question concerned which of four candidates is best for a job). In the Appendix I model this scenario and use the experimental probabilities for accuracy reported by Schulz-Hardt *et al.* (2006) to show that we can get a free riding scenario. This also helps to support my claim that this is a realistic worry, and not a merely possible one. Second, I have here only allowed for three kinds of beliefs to be adopted in any proposition: $c(P) = 1$, $c(P) = 0.5$, or $c(P) = 0$. This makes the calculations much simpler, but again, the model permits consideration of cases where agents can adopt any value for their credences. Though I don't go through it in any detail, in the Appendix I gesture towards how one could approach such cases.

6 Significance of Epistemic Free Riding

In the introduction to this paper I mentioned the Independence Thesis (Mayo-Wilson *et al.*, 2011), which claims that rational individuals can form irrational groups and that rational groups might be composed of irrational individuals. In this section I'd like to consider this thesis more closely and its relation to the epistemic free riding scenario just presented.

One way that one might argue for a tension between group and individual rationality is by pointing out that what it is rational for me to believe given that I am working alone, is different than what it is rational for me to believe given that I am part of a group. But this claim is not surprising. Since large groups of people can investigate more topics than individuals, I should presumably form stronger beliefs on more varied topics when part of a group than when working alone. There isn't any real conflict between group and individual rationality here, any more than there is a conflict between rationality in situations where there is not much evidence and rationality in situations when there is lots of evidence.

In the scenario I have presented, in contrast, we are presented with a genuine conflict between the group and individual perspective. To see why, suppose we say that an individual is rational iff she minimizes her own expected inaccuracy over time and say that a group is rational iff the average expected inaccuracy of the group is minimized. I don't mean to endorse these definitions of group and individual rationality, but they will help to illustrate the sense in which the free riding scenario here vindicates a certain version of the Independence Thesis. So, granting those definitions, the free riding scenario above shows that rational individuals can make up an irrational group. When I'm in a group setting, it really is epistemically better for me (and every other group member) to withhold belief upon learning of our disagreement. So, from the perspective of individual rationality, each group member should withhold. However, if I'm focused purely on the group performance I should prefer that we all remain steadfast in the face of disagreement. So, the free riding scenario suggests that rational groups will contain irrational members.[21] Thus, what is rational for an individual *qua* group member is different than what is rational for an individual *qua* individual.

Note further that this is a purely epistemic vindication of the Independence Thesis. Many have argued for a tension between group and individual rationality by focusing on the way that communication networks between researchers are structured, or the projects that different group members choose to spend their time pursuing.[22] While these topics are clearly im-

[21]The free riding scenario doesn't quite show that rational groups contain irrational members, since I haven't said anything about which kind of group will minimize expected accuracy, but it does suggest that whichever group it is, it will contain irrational members.

[22]See, for instance, Zollman (2007), Muldoon & Weisberg (2009), and Mayo-Wilson *et al.* (2011).

portant for the philosophy of science and relevant to epistemology, they are not purely about the belief states adopted by researchers. Thus, these discussions of the Independence Thesis run the risk of being characterized as not genuinely about *epistemic* rationality. Not so here. Here we are focused purely on the epistemic acts of adopting various belief states, which are then evaluated solely in terms of accuracy. It is thus a vindication of a purely epistemic Independence Thesis.

7 A Puzzle

I've argued that being conciliatory in group inquiry when one learns that there is disagreement is often to be an epistemic free rider. That is, it is in each person's epistemic self-interest to be conciliatory, even though the effect of everyone acting in this way is epistemically detrimental for everyone.

Against this backdrop, it is interesting to note that there is evidence that humans are naturally disposed to stand their ground in disagreements and not be conciliatory. Minson *et al.* (unpublished), for instance, give considerable evidence for what they call *disagreement neglect* – the phenomenon where disagreeing parties ignore disagreement and simply stick to their own views. Summarizing their results and the results of other researchers, they write:

> Discounting of peer input has been demonstrated with American, Israeli, French, and multi-national samples. In our study, both American and Japanese participants as well as ingroup and outgroup dyad members gave roughly twice as much weight to their own judgments as those of their partner, suggesting that the phenomenon is indeed a feature of basic human judgment. (p. 16)

It is worth noting that this goes against a claim that Thomas Kelly (2005) makes in an early paper on peer disagreement. In that paper Kelly claims that the empirical evidence suggests that most people are conciliatory in cases of disagreement: "There is a considerable amount of empirical evidence which suggests that an awareness of disagreement tends to lead us to significantly moderate our opinions." (pp. 170-1). In support of this, Kelly cites the classic studies by Solomon Asch from the 1950s.[23] These studies involve situations where there is a group of people, and all the group members are posed the same simple question, which they each must answer,

[23]See, for instance, Asch (1955, 1956).

sequentially and in front of the other group members. For instance, the question might be which two of three lines on the board are equal length. In the experiment, all the group members are confederates of the experimenter except one. The confederates all give a clearly incorrect answer to the question. Asch's results show that this can have a strong influence on the answer that the experimental subject gives. Though Kelly is correct that this does reveal a certain kind of conciliatory tendency, it is not strong evidence that in all cases of disagreement, humans have such a tendency. Instead, this shows a kind of conciliatory behavior when one is strongly outnumbered in a group situation. However, as Minson *et al.* (unpublished) note, we see a kind of non-conciliatory behavior when the groups are more evenly split. In fact, some of Asch's own modifications of his experiments show this. For instance, when Asch instructs just one of the confederates to give the correct answer, the experimental subject is much less likely to go along with majority (and obviously incorrect) opinion.[24] Thus, in the cases of interest—the cases where there is a roughly even split of opinion and so conciliatory attitudes are plausibly supported by the evidence—humans appear to be naturally non-conciliatory.

This mirrors, in an interesting way, what we see in the practical domain. There are many practical opportunities to free ride, to take an action that is in one's own self-interest even though if everyone took that action it would be bad for everyone. Nevertheless, it is commonplace that most of us are not disposed to act in our self-interest in these cases. Consider an example. Most of us recognize that when we are in a national park there is no self-interested reason not to just leave our garbage in the park, rather than carrying it out. Nevertheless, most of us are also disposed not to just leave our garbage in the park, perhaps out of some notion of fairness or responsibility. So there's a puzzle here: why are we disposed to act against our own self-interest in such cases? The case of epistemic free riding in group inquiry presents us with an analogous puzzle.

In the practical domain, different hypotheses have been given to try to explain this puzzle. Some of the more interesting hypotheses are evolutionary. According to these hypotheses, we are disposed to not be free riders because of some evolutionary advantage obtained by such a disposition. There are more and less controversial versions of this hypothesis.

The more controversial versions posit group-level selection. According to this view, a strategy that is disadvantageous to the individual may nevertheless be selected for in a population since it renders that population

[24] See, for instance, Asch (1955).

stronger than a population that plays the free riding strategy. Though intuitively clear, such a view has trouble explaining why such non-free-riding populations are not overrun by a mutant who happens to play the free riding strategy. If such a strategy really is beneficial to the individual, it should eventually win out in a population.

The less controversial versions do not posit group-level selection. Brian Skyrms (1996) develops this kind of idea in the framework of evolutionary game theory. In Skyrms's model the selection is not taking place at the group level. Rather, each individual is modeled as reproducing in proportion to the resources she herself gathers. Skyrms shows that under a wide variety of different strategies that those in your own population might be playing, those who play the free riding strategy will tend to die out. This is primarily due to the fact that many of the alternative strategies do not engage in cooperative behavior with free riders. Here although individual instances of free riding may be beneficial to the individual, the free riding disposition is not, in general, one that survives.

One possible explanation for how we avoid the problem of epistemic free riding, then, models itself on these evolutionary hypotheses: perhaps there is some evolutionary pressure towards the disposition to remain steadfast rather than to be conciliatory in cases of disagreement. This would fit well with the fact that we do appear to have such a disposition.

It is possible to give a plausible-sounding group-level selection story:

> Groups that are quickly conciliatory in cases of disagreement are not as good at finding correct answers to questions of importance. Thus, groups with such conciliatory tendencies in the past would have been less successful in answering important questions related to survival such as where to plant crops, where to hunt, and how many resources to stockpile for difficult times. This would lead groups who were non-conciliatory to have an evolutionary advantage over groups with conciliatory tendencies. We are thus naturally disposed to stick to our initial views and defend them because such a disposition keeps most of us from being epistemic free riders and thus allows us to each play our part in group inquiry.

Unfortunately, this story depends on the controversial idea of group-level selection.

It is harder to see how a Skyrms-like story could be told for the disposition to remain steadfast in cases of disagreement. What we would need is some evidence that those who are disposed to be conciliatory are given less opportunity to cooperate with others in inquiry, and so end up gathering less accurate beliefs over the long haul despite the fact that in any one situation they can do better than those who are steadfast. It is hard, however, to see how to make such a story plausible.

8 Conclusion

In this paper I've argued that there are realistic cases of group inquiry where the problem of epistemic free riding can arise. These cases are structurally analogous to the more familiar cases of practical free riding.

Should we care about such cases? I think we should. First, the existence of such cases adds to the growing literature suggesting that group inquiry may in some cases be fundamentally different than individual inquiry. Second, cases of epistemic free riding show that even pure inquirers—those who care only about finding the truth—may sometimes have overriding reason to believe propositions that they are not justified in believing. Third, the fact that being conciliatory in a debate is often an instance of free riding raises an interesting question about why humans nevertheless appear to be naturally disposed to being steadfast in the face of controversy.

9 Appendix

9.1 More Than Two Possible Answers

In the main text is a model of a scenario where a group is attempting to answer a question with two possible answers. In the empirical study by Schulz-Hardt *et al.* (2006), there were four possible answers to the questions the groups attempted to solve. The homogenous groups were correct 7% of the time, whereas the heterogenous groups were correct 43% of the time. Though there are some things about the study that make it not perfectly translatable to our model, it'd be nice to see if in using these probabilities we get an epistemic free riding scenario.[25]

Assume, as before that agents can either remain steadfast in their prediscussion opinions or withhold belief. In this case, however, there are four possible answers to the question. Accordingly, we can model the option

[25]What features make the study not perfectly translatable? Given the experimental design, the homogenous groups never had a group member who initially prefers the correct answer. Further, the homogenous groups all prefer the same one answer to the question; they are not homogenous in withholding belief.

of withholding belief as assigning credence of 1/4 to each of these possible answers. We will also require that groups do not withhold belief at the end of inquiry since this follows the experimental design of Schulz-Hardt *et al.* (2006) where groups are forced to settle on an answer. Given this, $t_{GC} = 0.07$, $f_{GC} = 0.93$, $t_{GS} = 0.43$, and $f_{GS} = 0.57$. Suppose we stick with the Brier Score. Then we have:

$$a = EB(C) + EB(GC) = 0.75 + 1.86 = 2.61$$

$$d = EB(S) + EB(GS) = 1.5 + 1.14 = 2.64$$

So, in this particular case, $d > a$ and the Pareto condition is not met.

However, this result depends crucially on the scoring rule used. For instance, if we use a different proper scoring rule, the spherical score, we get a different verdict. The spherical score is given by:

$$SS(\mathbf{c}, i) = \frac{c_i}{(\sum_{j=1}^{m} c_j^2)^{1/2}}.$$

It has a minimum value of 0 and a maximum value of 1, independent of m. Unlike the Brier Score, however, higher numbers are better. For continuity with the Brier Score, then, let us work with a simple transformation: $SS^*(\mathbf{c}, i) = 1 - SS(\mathbf{c}, i)$. SS^* can thus also be seen as measuring inaccuracy. Using this we have:

$$a = ESS^*(C) + ESS^*(GC) = 0.5 + 0.93 = 1.43$$

$$d = ESS^*(S) + ESS^*(GS) = 0.75 + 0.57 = 1.32$$

In this case $d < a$ and so the Pareto condition is met. This suggests that the scenario investigated by Schulz-Hardt *et al.* (2006) is a borderline case of a free riding scenario. According to some plausible scoring rules, it is a case of free riding; according to others, it isn't. Altogether, however, this strengthens the case for the claim that there are realistic cases where the conditions will be met for epistemic free riding.

Explanation of values:
Let the possible answers to the question be in $\Omega = \{1, 2, 3, 4\}$.

- $EB(C) = \sum_{i=1}^{4} p(i) \times B(c^W, i)$. For each value of i, $B(c^W, i) = (1 - 1/4)^2 + 3 \times (0 - 1/4)^2 = 0.75$. From this it follows that $EB(C) = 0.75$.

- $EB(S) = \sum_{i=1}^{4} p(i) \times B(c^n, i)$, where c^n assigns credence 1 to one of the possible answers in Ω and 0 to the others. When $i = n$, $B(c^n, i) = (1-1)^2 + 3 \times (0-0)^2 = 0$. When $i \neq n$, $B(c^n, i) = (1-0)^2 + (0-1)^2 + 2 \times (0-0)^2 = 2$. Since $p(i) = 1/4$, it follows that $EB(S) = 1.5$.

- $ESS^*(C) = \sum_{i=1}^{4} p(i) \times SS^*(c^W, i)$. For each value of i, $SS^*(c^W, i) = 1 - \frac{1/4}{(4 \times 1/4^2)^{1/2}} = 0.5$. From this it follows that $ESS^*(C) = 0.5$.

- $ESS^*(S) = \sum_{i=1}^{4} p(i) \times SS^*(c^n, i)$, where c^n assigns credence 1 to one of the possible answers in Ω and 0 to the others. When $i = n$, $SS^*(c^n, i) = 1 - \frac{1}{(1^2+3\times 0^2)^{1/2}} = 0$. When $i \neq n$, $SS^*(c^n, i) = 1 - \frac{0}{(1^2+3\times 0^2)^{1/2}} = 1$. Since $p(i) = 1/4$, it follows that $ESS^*(S) = 0.75$.

- $EB(GC)$ can be worked out simply by using t_{GC} and f_{GC}, the probability that a conciliatory group gets it right or wrong at the end of inquiry. Since in this simple model getting it right corresponds to assigning the true answer credence 1, we know that the Brier score for getting it right is $(1-1)^2+3\times(0-0)^2 = 0$. Similarly, since getting it wrong corresponds to assigning one of the false answers credence 1, we know that the Brier score for getting it wrong is $(1-0)^2+(0-1)^2+2\times(0-0)^2 = 2$. Accordingly, $EB(GC) = t_{GC} \times 0 + f_{GC} \times 2 = 0.93 \times 2 = 1.86$.

- $EB(GS)$ is worked out similarly. The only difference is in the probabilities of getting it right or wrong. Thus, $EB(GS) = t_{GS} \times 0 + f_{GS} \times 2 = 0.57 \times 2 = 1.14$.

- Similar calculations show that $ESS^*(GC) = 0.93$ and that $ESS^*(GS) = 0.57$.

9.2 Complex Belief States

The model described in the text allows only three kinds of belief states. It is relatively straightforward, however, to generalize the model to allow for beliefs that take any value in $[0, 1]$. The expected Brier score for an epistemic act given in the text is:

$$EB(X) = \sum_i \sum_n p(i \wedge \mathbf{c}^n | X) \times B(\mathbf{c}^n, i)$$

To adjust this to allow any degree of belief to be adopted, there are two key moves. First, we must integrate rather than sum over the possible belief states, so n will no longer be finite. Second, we must define a probability distribution for $p(i \wedge \mathbf{c}^n|X)$.

This second step can be tricky if we allow questions with more than one answer. In that case, $\mathbf{c}^n$ will be a probability vector with more than two members. However, if we restrict our attention to cases where there are only two members, things are easier. I'll only explore this easier case here. One natural way to approach the probability distribution is to use a normal distribution. For instance, let $f_{\mu,\sigma}(x)$ be a normal distribution with mean μ and standard deviation σ. Let us let us interpret '$\mathbf{c}^n$' as referring to the belief state that assigns answer 1 a credence of n. Since there are only two possible answers, there is only one such belief state. Then, set $p(1 \wedge \mathbf{c}^n|X) = f_{\mu,\sigma}(n)/(2\int_0^1 f_{\mu,\sigma}(x)dx)$, and $p(2 \wedge \mathbf{c}^n|X) = f_{\mu,\sigma}(1-n)/(2\int_0^1 f_{\mu,\sigma}(x)dx)$, for appropriate μ and σ according to the epistemic act X.[26]

This yields the following:

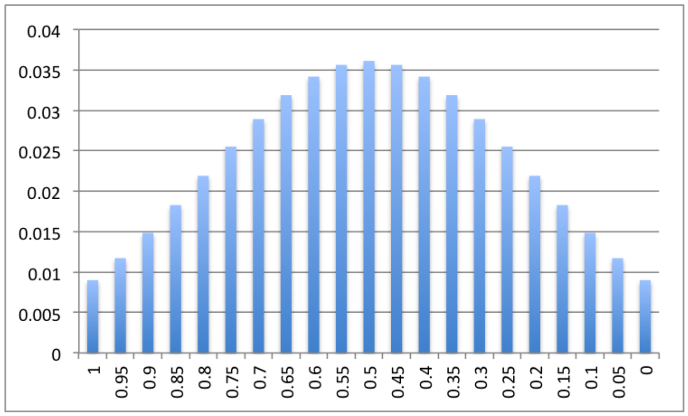
$$EB(X) = \sum_i \int_{n=0}^{n=1} p(i \wedge \mathbf{c}^n|X) \times S(\mathbf{c}^n, i),$$

with $p(i \wedge \mathbf{c}^n|X)$ defined as described immediately above using a normal distribution.

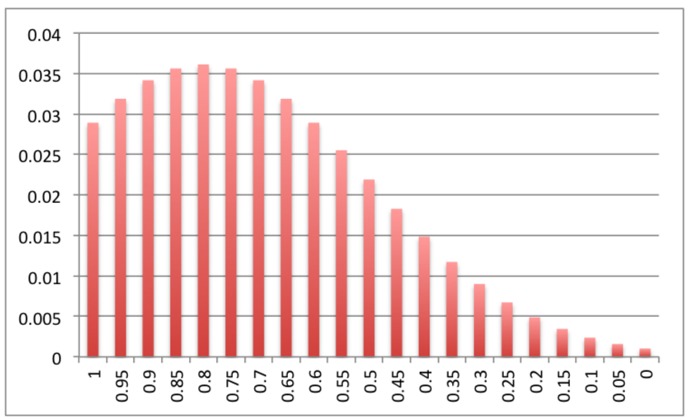
What *is* an appropriate value for μ and σ? In virtue of the empirical data, μ should be greater for the option of remaining steadfast than for the option of being conciliatory. This corresponds to it being more likely that the steadfast group gets closer to the truth. To get an intuitive picture of the idea, here is one specific way to model this:

[26] The integral in the numerator is simply to normalize the values. It is multiplied by 2 so that $\int_{n=0}^{n=1} p(1 \wedge \mathbf{c}^n|X) + \int_{n=0}^{n=1} p(2 \wedge \mathbf{c}^n|X) = 1$.

Conciliatory:



Steadfast:



On the x-axis is the credence assigned to answer 1 when it is true. On the y-axis is the probability of this scenario obtaining. (These charts show only the values of $p(1 \wedge \mathbf{c}^n|X)$.) The conciliatory distribution has a mean of 0.5 while the steadfast distribution has a mean of 0.8. This corresponds to the conciliatory group being less likely to get close to the truth than the steadfast group.

References

Asch, S. (1955). Opinions and social pressure. *Scientific American*, *193*(5), 31–35.

Asch, S. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*(9), 1–70.

Banerjee, A. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, *107*(3), 797–817.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, *100*(5), 992–1026.

Christensen, D. (2007). Epistemology of disagreement: the good news. *Philosophical Review*, *116*, 187–217.

Dunn, J. (ms). Peer disagreement and group inquiry.

Elga, A. (2007). Reflection and disagreement. *Noûs*, *41*(3), 478–502.

Evans, J. S. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*(3), 295–306.

Gigone, D. & Hastie, R. (1997). Proper analysis of the accuracy of group judgments. *Psychological Bulletin*, *121*(1), 149–167.

Greaves, H. (2013). Epistemic decision theory. *Mind*, *122*, 915–952.

Greitemeyer, T., Schulz-Hardt, S., Brodbeck, F. C., & Frey, D. (2006). Information sampling and group decision making: The effects of an advocacy decision procedure and task experience. *Journal of Experimental Psychology: Applied*, *12*(1), 31–42.

Hastie, R. (1986). Experimental evidence on group accuracy. In B. Grofman & G. Owen (Eds.), *Information Pooling and Group Decision Making: Proceedings of the Second University of California, Irvine Conference on Political Economy*, JAI Press. 129–157.

Hawthorne, J. A. (ms). Voting in search of the public good: the probabilistic logic of majority judgments. Available at: http://faculty-staff.ou.edu/H/James.A.Hawthorne-1/Hawthorne–Jury-Theorems.pdf.

Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science*, *65*(4), 575–603.

Kelly, T. (2005). The epistemic significance of disagreement. *Oxford studies in epistemology*, *1*, 167–196.

Kitcher, P. (1990). The division of cognitive labor. *The Journal of Philosophy*, *87*(1), 5–22.

Konek, J. & Levinstein, B. (ms). The foundations of epistmemic decision theory. http://jasonkonek.com/FEUT.pdf.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentive reasoning. *Cognition and Instruction*, *15*(3), 287–315.

List, C. & Pettit, P. (2004). An epistemic free riding problem? In P. Catton & G. Macdonald (Eds.), *Karl Popper: Critical Appraisals*, Routledge. 128–158.

Mayo-Wilson, C., Zollman, K. J., & Danks, D. (2011). The independence thesis: When individual and social epistemology diverge. *Philosophy of Science*, *78*(4), 653–677.

Mercier, H. (2011). Reasoning serves argumentation in children. *Cognitive Development*, *26*(3), 177–191.

Mercier, H. (2012). Looking for arguments. *Argumentation*, *26*(3), 305–324.

Mercier, H. & Sperber, D. (2011). Why do humans reason? arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*, 57–111.

Minson, J. A., Mercier, H., Deguchi, M., & Yama, H. (unpublished). Disagreement neglect: Failure to benefit from divergent judgments is moderated by group belonging and culture. Available at: http://opim.wharton.upenn.edu/∼jminson/Site/PAPERS_files/Minson_et_al_2012.pdf.

Moshman, D. & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Educational Psychology Papers and Publications*, 52.

Muldoon, R. (2013). Diversity and the division of cognitive labor. *Philosophy Compass*, *8*(2), 117–125.

Muldoon, R. & Weisberg, M. (2009). Epistemic landscapes and the division of cognitive labor. *Philosophy of Science*, *76*(2), 225–252.

Perret-Clermont, A.-N., Carugati, F., & Oates, J. (2004). A sociocognitive perspective on learning and cognitive development. In J. Oates & A. Grayson (Eds.), *Cognitive and language development in children*, Wiley-Blackwell, chap. 8.

Pettit, P. (1986). Free riding and foul dealing. *The Journal of Philosophy*, 361–379.

Schulz-Hardt, S., Brodbeck, F., Mojzisch, A., Kerschreiter, R., & Frey, D. (2006). Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology*, *91*(6), 1080–1093.

Schulz-Hardt, S., Jochims, M., & Frey, D. (2002). Productive conflict in group decision making: genuine and contrived dissent as strategies to counteract biased information seeking. *Organizational Behavior and Human Decision Processes*, *88*(2), 563–586.

Seidenfeld, T. (1985). Calibration, coherence, and scoring rules. *Philosophy of Science*, *52*, 274–294.

Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge University Press.

Strauss, S. G., Parker, A. M., & Bruce, J. B. (2011). The group matters: A review of processes and outcomes in intelligence analysis. *Group Dynamics: Theory, Research, and Practice*, *15*(2), 128–146.

Strevens, M. (2003). The role of the priority rule in science. *Journal of Philosophy*, *100*(2), 55–79.

Sunstein, C. (2002). The law of group polarization. *Journal of Political Philosophy*, *10*(2), 175–195.

Wedgewood, R. (ms). The probabilistic structure of rational entitlement. `http://www-bcf.usc.edu/ wedgwood/Entitlement.pdf`.

Zollman, K. J. (2007). The communication structure of epistemic communities. *Philosophy of Science*, *74*(5), 574–587.

Zollman, K. J. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, *72*(1), 17–35.